Encoding of Functions of Correlated Sources*

Saravanan Vijayakumaran†

February 1, 2008

Abstract

In this correspondence, we describe the achievable rate region for reliably recovering deterministic functions of correlated sources which have a finite alphabet. The method of proof is almost the same as that used to prove the Slepian-Wolf theorem.

1 Introduction

Consider the problem of recovering a function $F(X, Y)$ of two correlated sources (X, Y) by encoding the sources separately (see Fig. 1.) A problem of this class was first considered in [1], where the exact rate region for the modulo-two adder source network was derived. In [2], necessary and sufficient conditions were derived, for the achievable rate region for recovering functions of correlated sources to coincide with the Slepian-Wolf region [3].

In this correspondence, we describe the achievable rate region for reliably recovering deterministic functions of correlated sources which have a finite alphabet. The method of proof is almost the same as that used to prove the Slepian-Wolf theorem [3], [4].

2 System Model

The system model is essentially the same as the one described in [2]. We repeat it here for convenience and notational clarity.

Let X and Y be a pair of correlated random variables defined on finite sample spaces $\mathscr{X}$ and $\mathscr{Y}$, respectively. Denote their joint probability distribution by

$$p_{XY}(x, y) = \Pr[X = x, Y = y], \qquad x \in \mathscr{X}, y \in \mathscr{Y}. \tag{1}$$

Conforming with the usual convention, we will use uppercase letters to denote random variables and lowercase letters to denote fixed values the random variables may take. Let $(\mathbf{X}, \mathbf{Y}) = (X^n, Y^n) = ((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ be a sequence of n independent realizations of the pair of random variables (X, Y). The distribution of $(\mathbf{X}, \mathbf{Y})$ is given by

$$p_{XY}(\mathbf{x}, \mathbf{y}) = \Pr[\mathbf{X} = \mathbf{x}, \mathbf{Y} = \mathbf{y}] = \prod_{i=1}^{n} p_{XY}(x_i, y_i), \qquad \mathbf{x} \in \mathscr{X}^n, \mathbf{y} \in \mathscr{Y}^n. \tag{2}$$

The number of coordinates in $(\mathbf{X}, \mathbf{Y})$ or $(\mathbf{x}, \mathbf{y})$ will be clear from context.

*This work was supported in part by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322, and by a University of Florida Alumni Fellowship (2001-2005).

†The author is with the Laboratory for Computer Communications and Applications, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

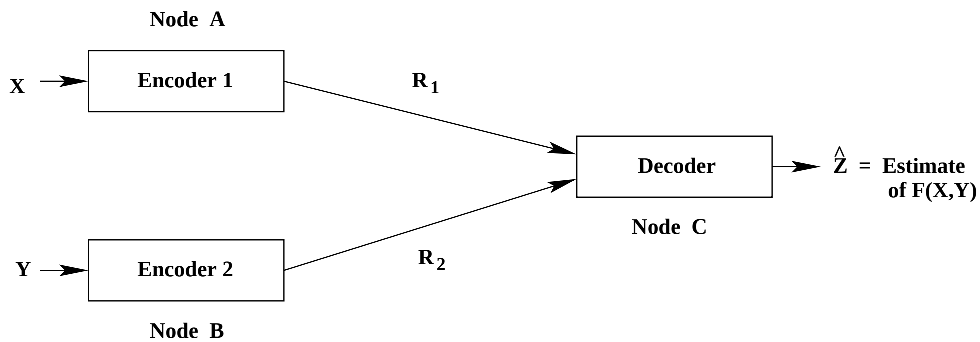

Figure 1: Illustration of the system model.

Let $F : \mathscr{X} \times \mathscr{Y} \mapsto \mathscr{Z}$ be an arbitrary deterministic function. We will denote the sequence $(F(X_1, Y_1), F(X_2, Y_2), \ldots, F(X_n, Y_n))$ by $F(\mathbf{X}, \mathbf{Y})$. We will sometimes find it convenient to denote the random variable $F(X, Y)$ by Z. Then $\mathbf{Z} = Z^n = F(\mathbf{X}, \mathbf{Y})$.

The sequence $(X_1, X_2, \ldots)$ is available at node A and the sequence $(Y_1, Y_2, \ldots)$ is available at node B. We wish to reliably recover the sequence $(Z_1, Z_2, \ldots)$ at node C, under the condition that there is no communication between nodes A and B. This situation is illustrated in Fig. 1. The channels from node A to node B and node A to node C are assumed to be noiseless. So we have a distributed source coding problem where the goal is to simultaneously minimize the required rates R_1 and R_2, which allow reliable recovery of the sequence $(Z_1, Z_2, \ldots)$ at node C.

We now present some definitions similar to ones presented in [4, Section 14.4].

Definition: A distributed source code $\mathscr{C}_n(F)$ for the random variable $F(X, Y)$ is a triplet of functions (f_1, f_2, g),

$$\begin{aligned} f_1 &: \mathscr{X}^n \mapsto \{1, 2, \ldots, 2^{nR_1}\} \\ f_2 &: \mathscr{Y}^n \mapsto \{1, 2, \ldots, 2^{nR_2}\} \\ g &: \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_2}\} \mapsto \mathscr{Z}^n \end{aligned}$$

where f_1, f_2 correspond to the encoding functions and g corresponds to the decoding function. Here R_1, R_2 are nonnegative real numbers and n is a positive integer.

Definition: For a particular distributed source code $\mathscr{C}_n(F)$, the probability of error is defined as

$$P_e^{(n)} = \Pr[g(f_1(\mathbf{X}), f_2(\mathbf{Y})) \neq F(\mathbf{X}, \mathbf{Y})]. \tag{3}$$

Definition: A rate pair (R_1, R_2) is said to be achievable for a function F if there exists a sequence of distributed source codes $\{\mathscr{C}_n(F) : n \in \mathbb{N}\}$ with corresponding probabilities of error $P_e^{(n)}$ such that $P_e^{(n)} \to 0$ as $n \to \infty$.

Definition: For a particular function F, the achievable rate region $\mathscr{R}(F)$ is the closure of the set of all achievable rate pairs.

3 Main Result

The following is the main result of this correspondence.

Theorem: The achievable rate region for a function F of correlated random variables (X, Y) is given by

$$\mathscr{R}(F) = \{(R_1, R_2) : R_1 \geq H(F(X,Y)|Y), R_2 \geq H(F(X,Y)|X), R_1 + R_2 \geq H(F(X,Y))\}.$$

The proof of this result is a simple application of the techniques used to prove the Slepian-Wolf theorem in [4]. So we shamelessly adopt the conventions and notation in [4, Chapter 14], if not for any other reason but to illustrate the simplicity of the proof. We need to borrow the following notation[1] before we proceed with the proof.

Let $(U_1, U_2, \ldots, U_k)$ be a finite collection of discrete random variables with a fixed joint distribution, $p(u_1, u_2, \ldots, u_n)$, $(u_1, u_2, \ldots, u_n) \in \mathscr{U}_1 \times \mathscr{U}_2 \times \cdots \times \mathscr{U}_k$. The set of ϵ-typical n-sequences will be denoted by $A_\epsilon^{(n)}(U_1, U_2, \ldots, U_k)$. We will denote the set of U_i n-sequences that are jointly ϵ-typical with a particular U_j n-sequence, $\mathbf{u}_j$, by $A_\epsilon^{(n)}(U_i|\mathbf{u}_j)$.

Proof of Achievability : For each $\mathbf{x} \in \mathscr{X}^n$, set $f_1(\mathbf{x})$ to a value chosen from the set $\{1, 2, \ldots, 2^{nR_1}\}$ according to a uniform distribution. Similarly, for each $\mathbf{y} \in \mathscr{Y}^n$ set $f_2(\mathbf{y})$ to a value chosen from the set $\{1, 2, \ldots, 2^{nR_2}\}$ according to a uniform distribution. The encoding functions are revealed to the corresponding encoder and the decoder, i.e., the decoder needs to know both f_1 and f_2 while encoder i needs to know only f_i, $i = 1, 2$.

The encoding operation consists of encoder 1 and encoder 2 sending the values of $f_1(\mathbf{X})$ and $f_2(\mathbf{Y})$, respectively, to the decoder. Given the encoder outputs $(f_1(\mathbf{X}), f_2(\mathbf{Y})) = (i_0, j_0)$, the decoder outputs its estimate of $F(\mathbf{X}, \mathbf{Y})$, $\hat{\mathbf{Z}}$, to be $\mathbf{z}$ if there exists a unique sequence $\mathbf{z} \in \mathscr{Z}^n$ such that $(\mathbf{z}, \mathbf{x}, \mathbf{y}) \in A_\epsilon^{(n)}(Z, X, Y)$ for some $(\mathbf{x}, \mathbf{y}) \in \mathscr{X}^n \times \mathscr{Y}^n$ such that $f_1(\mathbf{x}) = i_0$ and $f_2(\mathbf{y}) = j_0$. Note that the pair $(\mathbf{x}, \mathbf{y})$ need not be unique.

The decoder operation is where the current coding scheme differs from Slepian-Wolf coding scheme. Of course, if F is the identity function, i.e., $F(x, y) = (x, y), \forall (x, y) \in \mathscr{X} \times \mathscr{Y}$, then the above decoder coincides with the decoder in the Slepian-Wolf coding scheme.

We now proceed with the analysis of the probability of error averaged over all possible encoder choices f_1, f_2. Let $E = \{\hat{\mathbf{Z}} \neq \mathbf{Z}\}$ denote the decoding error event. Then we have $E = E_0 \cup E_1 \cup E_2 \cup E_{12}$ where

$$\begin{aligned}
E_0 &= \left\{\exists \text{ no } \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{y}') \in A_\epsilon^{(n)}(Z, X, Y) \text{ for some } (\mathbf{x}', \mathbf{y}') \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), f_1(\mathbf{y}') = f_1(\mathbf{Y})\right\}, \\
E_1 &= \left\{\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{Y}) \in A_\epsilon^{(n)}(Z, X, Y) \text{ for some } \mathbf{x}' \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), \mathbf{z} = F(\mathbf{x}', \mathbf{Y}) \neq F(\mathbf{X}, \mathbf{Y})\right\}, \\
E_2 &= \left\{\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{X}, \mathbf{y}') \in A_\epsilon^{(n)}(Z, X, Y) \text{ for some } \mathbf{y}' \ni f_1(\mathbf{y}') = f_1(\mathbf{Y}), \mathbf{z} = F(\mathbf{X}, \mathbf{y}') \neq F(\mathbf{X}, \mathbf{Y})\right\}, \\
E_{12} &= \Big\{\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{y}') \in A_\epsilon^{(n)}(Z, X, Y) \text{ for some } (\mathbf{x}', \mathbf{y}') \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), f_1(\mathbf{y}') = f_1(\mathbf{Y}), \\
&\qquad\qquad \mathbf{z} = F(\mathbf{x}', \mathbf{y}') \neq F(\mathbf{X}, \mathbf{Y})\Big\}.
\end{aligned}$$

From the definition of jointly typical sequences [4], it is easy to see that

$$\Pr[E_0] \leq \Pr[(\mathbf{z}, \mathbf{x}', \mathbf{y}') \in \mathscr{X}^n \times \mathscr{Y}^n \times \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{y}') \notin A_\epsilon^{(n)}(Z, X, Y)] < \epsilon, \tag{4}$$

[1] See [4, Section 14.2] for definitions and properties.

for sufficiently large n. We bound $\Pr[E_1]$ in the following manner.

$$
\begin{aligned}
&\Pr[E_1] \\
&\quad = \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{Y}) \in A_\epsilon^{(n)}(Z,X,Y) \text{ for some } \mathbf{x}' \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), \mathbf{z} = F(\mathbf{x}', \mathbf{Y}) \neq F(\mathbf{X}, \mathbf{Y})] \\
&\quad \stackrel{(a)}{\leq} \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{Y}) \in A_\epsilon^{(n)}(Z,Y), \text{ for some } \mathbf{x}' \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), \mathbf{z} = F(\mathbf{x}', \mathbf{Y}) \neq F(\mathbf{X}, \mathbf{Y})] \\
&\quad = \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{y}) \in A_\epsilon^{(n)}(Z,Y) \text{ for some } \mathbf{x}' \ni f_1(\mathbf{x}') = f_1(\mathbf{x}), \mathbf{z} = F(\mathbf{x}', \mathbf{y}) \neq F(\mathbf{x}, \mathbf{y})] \\
&\quad \leq \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) \Pr[(\mathbf{z}, \mathbf{y}) \in A_\epsilon^{(n)}(Z,Y) : \text{For some } \mathbf{x}' \neq \mathbf{x}, f_1(\mathbf{x}') = f_1(\mathbf{x})] \\
&\quad \stackrel{(b)}{\leq} \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) 2^{-nR_1} |A_\epsilon^{(n)}(Z|\mathbf{y})| \\
&\quad \stackrel{(c)}{\leq} 2^{-nR_1} 2^{n(H(Z|Y)+2\epsilon)},
\end{aligned}
$$

where

(a) follows from the fact that for any $(\mathbf{z}, \mathbf{x}', \mathbf{y}) \in \mathscr{Z}^n \times \mathscr{X}^n \times \mathscr{Y}^n$, $(\mathbf{z}, \mathbf{x}', \mathbf{y}) \in A_\epsilon^{(n)}(Z,X,Y)$ $\Rightarrow (\mathbf{z}, \mathbf{y}) \in A_\epsilon^{(n)}(Z,Y)$,

(b) follows from the fact that we are averaging over all possible encoder choices for f_1 and the property that for a fixed $\mathbf{y} \in \mathscr{Y}^n$, $|\{\mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{y}) \in A_\epsilon^{(n)}(Z,Y)\}| = |A_\epsilon^{(n)}(Z|\mathbf{y})|$,

(c) follows from the fact that $|A_\epsilon^{(n)}(Z|\mathbf{y})| \leq 2^{n(H(Z|Y)+2\epsilon)}$ [4, Theorem 14.2.2].

The final bound on $\Pr[E_1]$ tends to zero as $n \to \infty$ if $R_1 > H(Z|Y) + 2\epsilon$. Thus for sufficiently large n, $\Pr[E_1] < \epsilon$. Similarly, we can show that $\Pr[E_2] < \epsilon$ for sufficiently large n if $R_2 > H(Z|X) + 2\epsilon$.

Note that $E_1 \subset E_{12}$ and $E_2 \subset E_{12}$. It then follows that $E = E_0 \cup E_1 \cup E_2 \cup E_{12} = E_0 \cup E_1 \cup E_2 \cup (E_{12} \cap E_1^c \cap E_2^c)$. We will find it easier to bound $E_{12} \cap E_1^c \cap E_2^c$ rather than bound E_{12} directly. We bound $\Pr[E_{12} \cap E_1^c \cap E_2^c]$ in the following manner.

$$
\begin{aligned}
&\Pr[E_{12} \cap E_1^c \cap E_2^c] \\
&\quad = \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : (\mathbf{z}, \mathbf{x}', \mathbf{y}') \in A_\epsilon^{(n)}(Z,X,Y) \text{ for some } \mathbf{x}' \neq \mathbf{X}, \mathbf{y}' \neq \mathbf{Y} \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), \\
&\qquad\qquad f_2(\mathbf{y}') = f_2(\mathbf{Y}), \mathbf{z} = F(\mathbf{x}', \mathbf{y}') \neq F(\mathbf{X}, \mathbf{Y})] \\
&\quad \stackrel{(a)}{\leq} \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : \mathbf{z} \in A_\epsilon^{(n)}(Z) \text{ for some } \mathbf{x}' \neq \mathbf{X}, \mathbf{y}' \neq \mathbf{Y} \ni f_1(\mathbf{x}') = f_1(\mathbf{X}), f_2(\mathbf{y}') = f_2(\mathbf{Y}), \\
&\qquad\qquad \mathbf{z} = F(\mathbf{x}', \mathbf{y}') \neq F(\mathbf{X}, \mathbf{Y})] \\
&\quad = \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) \Pr[\exists \mathbf{z} \in \mathscr{Z}^n : \mathbf{z} \in A_\epsilon^{(n)}(Z) \text{ for some } \mathbf{x}' \neq \mathbf{x}, \mathbf{y}' \neq \mathbf{y} \ni f_1(\mathbf{x}') = f_1(\mathbf{x}), f_2(\mathbf{y}') = f_2(\mathbf{y}), \\
&\qquad\qquad \mathbf{z} = F(\mathbf{x}', \mathbf{y}') \neq F(\mathbf{x}, \mathbf{y})] \\
&\quad \leq \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) \Pr[\mathbf{z} \in A_\epsilon^{(n)}(Z) : \text{For some } \mathbf{x}' \neq \mathbf{x}, \mathbf{y}' \neq \mathbf{y}, f_1(\mathbf{x}') = f_1(\mathbf{x}), f_2(\mathbf{y}') = f_2(\mathbf{y})] \\
&\quad \stackrel{(b)}{\leq} \sum_{\mathbf{x},\mathbf{y}} p_{XY}(\mathbf{x},\mathbf{y}) 2^{-nR_1} 2^{-nR_2} |A_\epsilon^{(n)}(Z)| \\
&\quad \stackrel{(c)}{\leq} 2^{-n(R_1+R_2)} 2^{n(H(Z)+\epsilon)},
\end{aligned}
$$

where

(a) follows from the fact that for any $(\mathbf{z}, \mathbf{x}', \mathbf{y}') \in \mathscr{Z}^n \times \mathscr{X}^n \times \mathscr{Y}^n$, $(\mathbf{z}, \mathbf{x}', \mathbf{y}') \in A_\epsilon^{(n)}(Z, X, Y)$ $\Rightarrow \mathbf{z} \in A_\epsilon^{(n)}(Z)$,

(b) follows from the fact that we are averaging over all possible encoder choices f_1, f_2 and from the definition of $A_\epsilon^{(n)}(Z)$,

(c) follows from the fact that $|A_\epsilon^{(n)}(Z)| \leq 2^{n(H(Z)+\epsilon)}$.

The final bound on $\Pr[E_{12} \cap E_1^c \cap E_2^c]$ can be made smaller than ϵ for sufficiently large n if $R_1 + R_2 > H(Z) + \epsilon$.

Thus, we have $\Pr[E] \leq \Pr[E_0] + \Pr[E_1] + \Pr[E_2] + \Pr[E_{12} \cap E_1^c \cap E_2^c] < 4\epsilon$ for sufficiently large n. Since the probability of error averaged over all codes is less than 4ϵ, there exists at least one code $\mathscr{C}_n^*(F)$ for which the average probability of error is less than 4ϵ. Since ϵ was arbitrary, we can construct a sequence of codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. The arbitrary choice of ϵ also implies that any rate pair (R_1, R_2) satisfying $R_1 > H(F(X,Y)|Y), R_2 > H(F(X,Y)|X), R_1 + R_2 > H(F(X,Y))$ is achievable. Since the achievable rate region is the closure of all achievable rates, we have

$$\mathscr{R}(F) \supset \{(R_1, R_2) : R_1 \geq H(F(X,Y)|Y), R_2 \geq H(F(X,Y)|X), R_1 + R_2 \geq H(F(X,Y))\}.$$

This completes the proof of the achievability. ■

Proof of Converse : This proof is once again very similar to the proof of the converse to the Slepian-Wolf theorem [4, Section 14.4.2].

Let (R_1, R_2) be an achievable rate pair. By definition, there exists a sequence of distributed source codes $\{\mathscr{C}_n(F) : n \in \mathbb{N}\}$ and hence a sequence of function triplets $\{(f_1^{(n)}, f_2^{(n)}, g^{(n)}) : n \in \mathbb{N}\}$, with $P_e^{(n)} = \Pr[g(f_1(\mathbf{X}), f_2(\mathbf{Y})) \neq F(\mathbf{X}, \mathbf{Y})]$ such that $P_e^{(n)} \to 0$ as $n \to \infty$.

For notational convenience, define $I_0^{(n)} = f_1^{(n)}(\mathbf{X})$ and $J_0^{(n)} = f_2^{(n)}(\mathbf{Y})$. By Fano's inequality, we have

$$\begin{aligned} H(F(\mathbf{X},\mathbf{Y})|I_0^{(n)}, J_0^{(n)}) &\leq P_e^{(n)} \log |\mathscr{Z}^n| + 1 \\ &= P_e^{(n)} n \log |\mathscr{Z}| + 1 = n\delta_n, \end{aligned} \tag{5}$$

where $\delta_n = P_e^{(n)} \log |\mathscr{Z}|$. We know that $\delta_n \to 0$ as $n \to \infty$. Since conditioning reduces entropy, we also have

$$H(F(\mathbf{X},\mathbf{Y})|\mathbf{Y}, I_0^{(n)}, J_0^{(n)}) \leq n\delta_n, \tag{6}$$

$$H(F(\mathbf{X},\mathbf{Y})|\mathbf{X}, I_0^{(n)}, J_0^{(n)}) \leq n\delta_n, \tag{7}$$

Following the notation in [4], we will write $U \to V \to W$ for some random variables U, V, W to mean that U and W are conditionally independent given V. For the problem under consideration, we have the following relations,

$$\begin{aligned} (I_0^{(n)}, J_0^{(n)}) &\to (\mathbf{X}, \mathbf{Y}) \to F(\mathbf{X}, \mathbf{Y}), \\ I_0^{(n)} &\to (\mathbf{X}, \mathbf{Y}) \to (F(\mathbf{X}, \mathbf{Y}), \mathbf{Y}), \\ J_0^{(n)} &\to (\mathbf{X}, \mathbf{Y}) \to (F(\mathbf{X}, \mathbf{Y}), \mathbf{X}). \end{aligned}$$

Application of the data processing inequality to each of the above relations and simple manipulations yield the following respective inequalities.

$$H(I_0^{(n)}, J_0^{(n)}|\mathbf{X}, \mathbf{Y}) \leq H(I_0^{(n)}, J_0^{(n)}|F(\mathbf{X}, \mathbf{Y})) \tag{8}$$

$$H(I_0^{(n)}|\mathbf{X}, \mathbf{Y}) \leq H(I_0^{(n)}|F(\mathbf{X}, \mathbf{Y}), \mathbf{Y}) \tag{9}$$

$$H(J_0^{(n)}|\mathbf{X}, \mathbf{Y}) \leq H(J_0^{(n)}|F(\mathbf{X}, \mathbf{Y}), \mathbf{X}) \tag{10}$$

Then we have a chain of inequalities

$$\begin{aligned}
n(R_1+R_2) &\geq H(I_0^{(n)},J_0^{(n)}) = I(F(\mathbf{X},\mathbf{Y});I_0^{(n)},J_0^{(n)}) + H(I_0^{(n)},J_0^{(n)}|F(\mathbf{X},\mathbf{Y}))\\
&\overset{(a)}{\geq} I(F(\mathbf{X},\mathbf{Y});I_0^{(n)},J_0^{(n)}) + H(I_0^{(n)},J_0^{(n)}|\mathbf{X},\mathbf{Y})\\
&\overset{(b)}{=} I(F(\mathbf{X},\mathbf{Y});I_0^{(n)},J_0^{(n)})\\
&= H(F(\mathbf{X},\mathbf{Y})) - H(F(\mathbf{X},\mathbf{Y})|I_0^{(n)},J_0^{(n)})\\
&\overset{(c)}{\geq} nH(F(X,Y)) - n\delta_n,
\end{aligned}$$

where

(a) follows from (8),

(b) follows from the fact that $(I_0^{(n)},J_0^{(n)})$ is a function of $(\mathbf{X},\mathbf{Y})$,

(c) follows from the chain rule and the fact that $F(\mathbf{X},\mathbf{Y})$ consists of i.i.d. components, and from (5).

Similarly, we can write

$$\begin{aligned}
nR_1 &\geq H(I_0^{(n)}) \geq H(I_0^{(n)}|\mathbf{Y})\\
&= I(F(\mathbf{X},\mathbf{Y});I_0^{(n)}|\mathbf{Y}) + H(I_0^{(n)}|F(\mathbf{X},\mathbf{Y}),\mathbf{Y})\\
&\overset{(a)}{\geq} I(F(\mathbf{X},\mathbf{Y});I_0^{(n)}|\mathbf{Y}) + H(I_0^{(n)}|\mathbf{X},\mathbf{Y})\\
&\overset{(b)}{=} I(F(\mathbf{X},\mathbf{Y});I_0^{(n)}|\mathbf{Y})\\
&= H(F(\mathbf{X},\mathbf{Y})|\mathbf{Y}) - H(F(\mathbf{X},\mathbf{Y})|\mathbf{Y},I_0^{(n)},J_0^{(n)})\\
&\overset{(c)}{\geq} nH(F(X,Y)|Y) - n\delta_n,
\end{aligned}$$

where

(a) follows from (9),

(b) follows from the fact that $I_0^{(n)}$ is a function of $\mathbf{X}$,

(c) follows from the chain rule and the fact that $H(F(X_i,Y_i)|Y_i) = H(F(X,Y)|Y)$ for $i = 1,2,\ldots,n$, and from (6).

Using similar techniques, we also get $nR_2 \geq nH(F(X,Y)|X) - n\delta_n$ by using (10) and (7). Thus, for any n, we have $R_1 \geq H(F(X,Y)|Y) - \delta_n$, $R_2 \geq H(F(X,Y)|X) - \delta_n$ and $R_1 + R_2 \geq H(F(X,Y)) - \delta_n$. Since $\delta_n \to 0$ as $n \to \infty$, we have that any rate pair is achievable only if $R_1 \geq H(F(X,Y)|Y)$, $R_2 \geq H(F(X,Y)|X)$ and $R_1 + R_2 \geq H(F(X,Y))$. Thus,

$$\mathscr{R}(F) \subset \{(R_1,R_2) : R_1 \geq H(F(X,Y)|Y), R_2 \geq H(F(X,Y)|X), R_1+R_2 \geq H(F(X,Y))\}.$$

This completes the proof of the converse. ■

4 Concluding Remarks

We have found the exact achievable rate region for the problem of reliably recovering a function of correlated sources by separate encoding of the sources. The proof turns out to be a simple plug-and-play of the techniques in [4]. It is obvious that the achievable rate region found here reduces to the Slepian-Wolf region when F is the identity function. Although less obvious, it is not difficult to see that the result derived in this correspondence conforms with the results of [1], [2].

References

[1] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inform. Theory*, vol. 25, pp. 219–221, Mar 1979.

[2] T. S. Han and K. Kobayashi, "A dichotomy of functions $F(X, Y)$ of correlated sources (X, Y) from the viewpoint of the achievable rate region," *IEEE Trans. Inform. Theory*, vol. 33, Jan 1987.

[3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, July 1973.

[4] T. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.